

<https://helda.helsinki.fi>

Systematic mapping of cancer cell target dependencies using high-throughput drug screening in triple-negative breast cancer

Wang, Tianduanyi

2020

Wang , T , Gautam , P , Rousu , J & Aittokallio , T 2020 , ' Systematic mapping of cancer cell target dependencies using high-throughput drug screening in triple-negative breast cancer ' , Computational and Structural Biotechnology Journal , vol. 18 , pp. 3819-3832 . <https://doi.org/10.1016/j.csbj.2020.11.001>

<http://hdl.handle.net/10138/325869>

<https://doi.org/10.1016/j.csbj.2020.11.001>

cc_by_nc_nd

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

journal homepage: www.elsevier.com/locate/csbj

Systematic mapping of cancer cell target dependencies using high-throughput drug screening in triple-negative breast cancer

Tianduanyi Wang^{a,b}, Prson Gautam^a, Juho Rousu^b, Tero Aittokallio^{a,b,c,d,*}

^a Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland

^b Helsinki Institute for Information Technology (HIIT), Department of Computer Science, Aalto University, Espoo, Finland

^c Institute for Cancer Research, Department of Cancer Genetics, Oslo University Hospital, Oslo, Norway

^d Centre for Biostatistics and Epidemiology (OCBE), Faculty of Medicine, University of Oslo, Oslo, Norway



ARTICLE INFO

Article history:

Received 30 July 2020

Received in revised form 23 October 2020

Accepted 1 November 2020

Available online 24 November 2020

Keywords:

Target addiction

Cancer vulnerability

Functional dependency

Gene essentiality

Drug screening

Multi-targeted treatments

ABSTRACT

While high-throughput drug screening offers possibilities to profile phenotypic responses of hundreds of compounds, elucidation of the cell context-specific mechanisms of drug action requires additional analyses. To that end, we developed a computational target deconvolution pipeline that identifies the key target dependencies based on collective drug response patterns in each cell line separately. The pipeline combines quantitative drug-cell line responses with drug-target interaction networks among both intended on- and potent off-targets to identify pharmaceutically actionable and selective therapeutic targets. To demonstrate its performance, the target deconvolution pipeline was applied to 310 small molecules tested on 20 genetically and phenotypically heterogeneous triple-negative breast cancer (TNBC) cell lines to identify cell line-specific target mechanisms in terms of cytotoxic and cytostatic drug target vulnerabilities. The functional essentiality of each protein target was quantified with a target addiction score (TAS), as a measure of dependency of the cell line on the therapeutic target. The target dependency profiling was shown to capture inhibitory information that is complementary to that obtained from the structure or sensitivity of the drugs. Comparison of the TAS profiles and gene essentiality scores from CRISPR-Cas9 knockout screens revealed that certain proteins with low gene essentiality showed high target addictions, suggesting that they might be functioning as protein groups, and therefore be resistant to single gene knock-out. The comparative analysis discovered protein groups of potential multi-target synthetic lethal interactions, for instance, among histone deacetylases (HDACs). Our integrated approach also recovered a number of well-established TNBC cell line-specific drivers and known TNBC therapeutic targets, such as HDACs and cyclin-dependent kinases (CDKs). The present work provides novel insights into druggable vulnerabilities for TNBC, and opportunities to identify multi-target synthetic lethal interactions for further studies.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

In phenotype-based drug discovery, high-throughput screening (HTS) is often used to profile the sensitivity or toxicity of hundreds of drug molecules across panels of cell lines or patient-derived samples to identify a few hit compounds with desired response profiles for further development [1,2]. Compared to the target-based drug discovery, which starts from an individual, disease-related protein, phenotype-based drug discovery assays profile the complex cellular system as a whole, with the aim to reveal

small molecules or other bioactive molecules causing specific cellular phenotypes [3]. However, the underlying cellular response patterns and target mechanisms of the screened drugs often remain unknown in HTS experiments. Therefore, target deconvolution approaches are needed to map the observed cellular responses of drugs to the underlying drug-target interaction networks, with the aim to dissect the cell context-specific mechanisms of the drugs' action in terms of their on- and off-target profiles [3]. Various experimental methods have been used to identify potential therapeutic targets, such as chemical proteomics, gene expression response profiling and CRISPR-Cas9 knockout screening [4]. However, there is a need for systematic and cost-effective target deconvolution by computational means.

* Corresponding author.

E-mail address: tero.aittokallio@helsinki.fi (T. Aittokallio).

Due to the wide polypharmacological effects of many drug classes, a single drug may have potency against multiple protein targets. For instance, protein kinase inhibitors are tractable in drug development and play a role in a wide range of diseases, such as cardiovascular disorders and cancers [5,6]. However, protein kinase domains share structural and sequence similarity, and most kinase inhibitors bind to conserved ATP-binding pockets, which leads to prevalent target promiscuity and polypharmacological effects. Therefore, kinase inhibitors in particular require multi-target deconvolution approaches. The non-intended targets, termed as off-targets, are one of the main reasons for adverse drug reactions [7,8]. However, in some cases, especially in complex diseases such as cancer, off-targeting effects of kinase inhibitors may also lead to increased therapeutic effects through multi-targeting several cancer survival pathways [9]. In particular combinations of oncogene co-inhibitions may lead to selective killing of cancer cells, a concept generally termed as synthetic lethality [10,11], which can be harnessed through polypharmacological effects [12]. However, systematic identification of synthetic lethal interactions among multiple targets has remained difficult due to complex and context-specific interaction networks.

To facilitate the phenotype-based and context-specific drug discovery process, we implemented a computational target deconvolution pipeline (see Fig. 1). High-throughput drug screening assays are first quantified using response metrics, such as drug sensitivity score (DSS) [13], which provide a robust summary of the drug dose-responses in each cell line [14,15]. Next, we developed a confidence ranking approach for the heterogeneous drug-target binding information extracted from public target databases, such as ChEMBL [16] and Drug Target Commons [17], to map the observed drug responses to most reliable and potent target activities. Target addition score (TAS) [18] is then calculated for each target and cell line separately, based on the drug-target interaction networks, to quantify the functional importance of each target protein for the particular cell line inhibition. To provide statistical support for these druggable target dependencies, we implemented an empirical inference approach that uses a permutation procedure to statistically test the significance of an observed TAS level by generating null distribution based on randomized drug-target interaction networks. The novel permutation testing approach makes it possible to compare response profiles from HTS experiments with different sizes of compound libraries.

We applied the target deconvolution pipeline to identify functionally important target proteins in 20 triple negative breast cancer (TNBC) cell lines. TNBC is a highly aggressive subtype of breast cancer which is devoid of the hormonal receptors and human epidermal growth factor 2 (HER2). TNBC also lacks effective pharmaceutically targetable genetic alterations. Even though there are targeted therapies, such as poly (ADP-ribose) polymerase (PARP) inhibitors, standard chemotherapy often remains the only drug treatment option for TNBC [19,20]. However, since TNBC patients often have worse outcomes for chemotherapy [21,22], there is an intensive search for more efficient and targeted therapies for TNBC [23,24]. TNBC is a highly heterogeneous breast cancer subtype, hence demands stratified treatment options. To aid TNBC therapeutics stratification, the target dependency profiles of the TNBC cell lines, calculated based on HTS of 310 small molecules, were further correlated with gene essentiality profiles from CRISPR-Cas9 knock-out experiments. Clustering of the therapeutic targets based on drug-protein interaction profiles revealed protein groups that are resistant to single-gene knockout, hence providing insights into the druggable vulnerabilities for TNBC, and opportunities to identify multi-target synthetic lethal interactions for further studies.

2. Results

We screened 20 TNBC cell lines using a drug library of 310 small molecules, which includes both approved drugs and investigational compounds (Fig. S1). Two types of cell-based *in vitro* assays were conducted to profile drug-induced cellular phenotypes in 2D assays, namely cell viability and cell toxicity assays (see Materials and Methods). DSS was calculated to quantify both drug-induced cell viability inhibition and cytotoxicity. The TNBC cell lines exhibited heterogeneous drug responses across 310 oncology compounds in terms of both viability inhibition and cytotoxicity phenotypes [25]. This alarming result calls for the cell line-specific target deconvolution to better understand the cell context-specific target mechanisms of the drugs and to identify effective and selective therapeutic targeting options for stratified TNBC treatments.

2.1. Systematic mapping of drug-target interactions for target deconvolution

Construction of reliable drug-target interaction networks is hindered by multiple bioactivities measured for the same drug-target pair, reported by different labs and studies using various experimental conditions. To construct systems-wide drug-target interaction networks for target deconvolution, we therefore summarized the most confident drug-target bioactivities based on confidence rankings of all the experimental bioactivity measurements between the 310 drugs and 632 targets retrieved from the DTC database (see Materials and Methods). The confidence ranking of the various bioactivity categories was based on the variability of replicated bioactivity measurements within each category, estimated by the median absolute deviation (MAD). For 4995 bioactivities (36%), with annotated standard type and assay format available, we obtained a detailed confidence ranking across both biochemical and cell/tissue-based assays (Table 1). The confidence ranking for the remaining 8790 bioactivities (64%), for which only known standard type annotation was available, remained generally very similar (Table S2).

This confidence ranking enables one to carry out comprehensive target deconvolution across wide libraries of drugs and targets by integration of bioactivity data from various target profiling assays, even in the absence of all the experimental factors of the assays affecting their performance. The median of bioactivities from the most confident bioassay categories was here considered as the summarized bioactivity for each single drug-target pair (see Materials and Methods). Fig. 2 shows selected examples of drug-target pairs with their original and summarized bioactivities. It was observed that under different types of assay and detection technologies, the bioactivities for the same drug-target pair can easily span across two orders of magnitude in their potency. Such heterogeneity in the original bioactivities hinders the target deconvolution process, whereas the summarized bioactivities based on the confidence rankings and relatively standardized bioactivity levels enable construction of more reliable drug-target interaction networks, and further target deconvolution statistical analysis based on the most confident drug-target interactions only.

Interestingly, the bioactivity summarization process revealed that while the intended targets of drugs often have higher activities across their multiple on- and off-targets, as was expected, there exists also several drugs for which the nominal target activities are actually weaker than the median activities among their off-targets (Fig. 3, upper panel). This suggests that polypharmacological and drug off-target effects play an important role in the drug-induced cell line response phenotypes, as was argued also in the recent study [11], and therefore these potent off-target

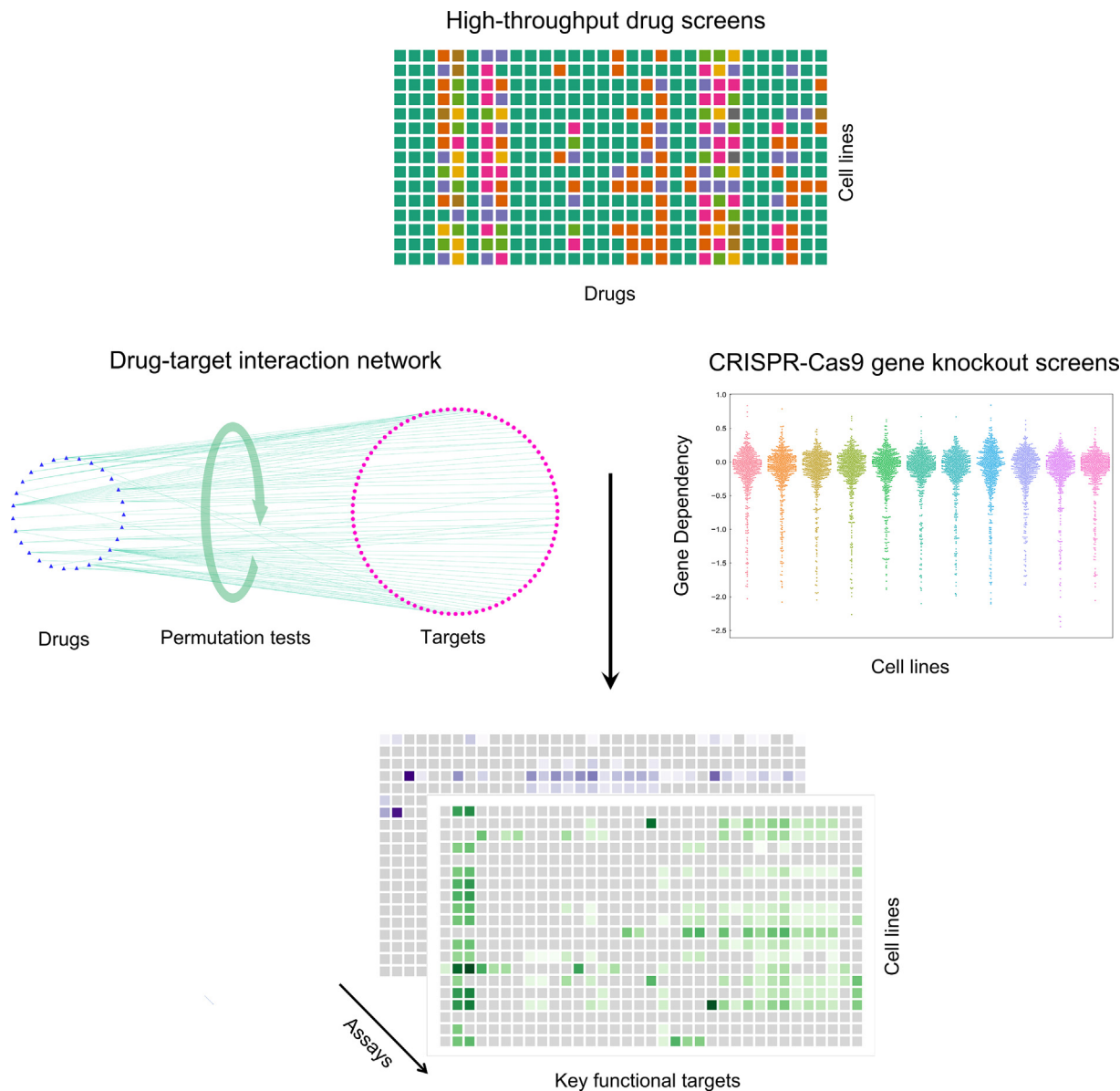


Fig. 1. Schematic overview of the target deconvolution workflow implemented in the study.

Table 1
Ranking of bioactivity assays based on median absolute deviation (MAD).

| Standard type | Assay format | Median MAD | Confidence Rank |
|---------------|----------------|------------|-----------------|
| K_i | Biochemical | 0,005 | 1 |
| K_d^{app} | Cell-free | 0,013 | 2 |
| K_d | Cell-based | 0,021 | 3 |
| K_d | Biochemical | 0,038 | 4 |
| IC_{50} | Cell-based | 0,064 | 5 |
| K_i | Cell-based | 0,074 | 6 |
| EC_{50} | Cell-based | 0,076 | 7 |
| IC_{50} | Biochemical | 0,079 | 8 |
| K_d | Physiochemical | 0,080 | 9 |
| IC_{50} | Tissue-based | 0,115 | 10 |
| IC_{50} | Cell-free | 0,126 | 11 |
| EC_{50} | Tissue-based | 0,235 | 12 |

effects needs to be taken into account in the target deconvolution process. In general, the bioactivity measurements using various multi-dose assay types (K_d , K_i , IC_{50} and EC_{50}) show relatively large variability also within the categories of both nominal on-targets

and non-intended off-targets (Fig. 3). This further motivates the use of an effective bioactivity summarization process that uses only the most confident bioactivity measurements when constructing drug-target networks for target deconvolution.

2.2. Characterization of the compound and target activity spaces of the screening library

To focus on potent compound-target activities, we filtered out all non-potent drug-target interactions and used only those pairs with a summarized bioactivity < 1000 nM in the target deconvolution. This activity filtering resulted in 4 276 unique bioactivities quantifying potent binding interactions between the 310 drugs and 632 targets. The distributions of bioactivities before and after the summarization and activity filtering show that the activity filtering process effectively eliminates the non-potent interactions, whereas the right-hand tail of active interactions has a relatively similar shape (Fig. 4A-B). When using the rather liberal activity cut-off of 1000 nM, the average number of targets per drug was 4 in our dataset, although there were also compounds with more

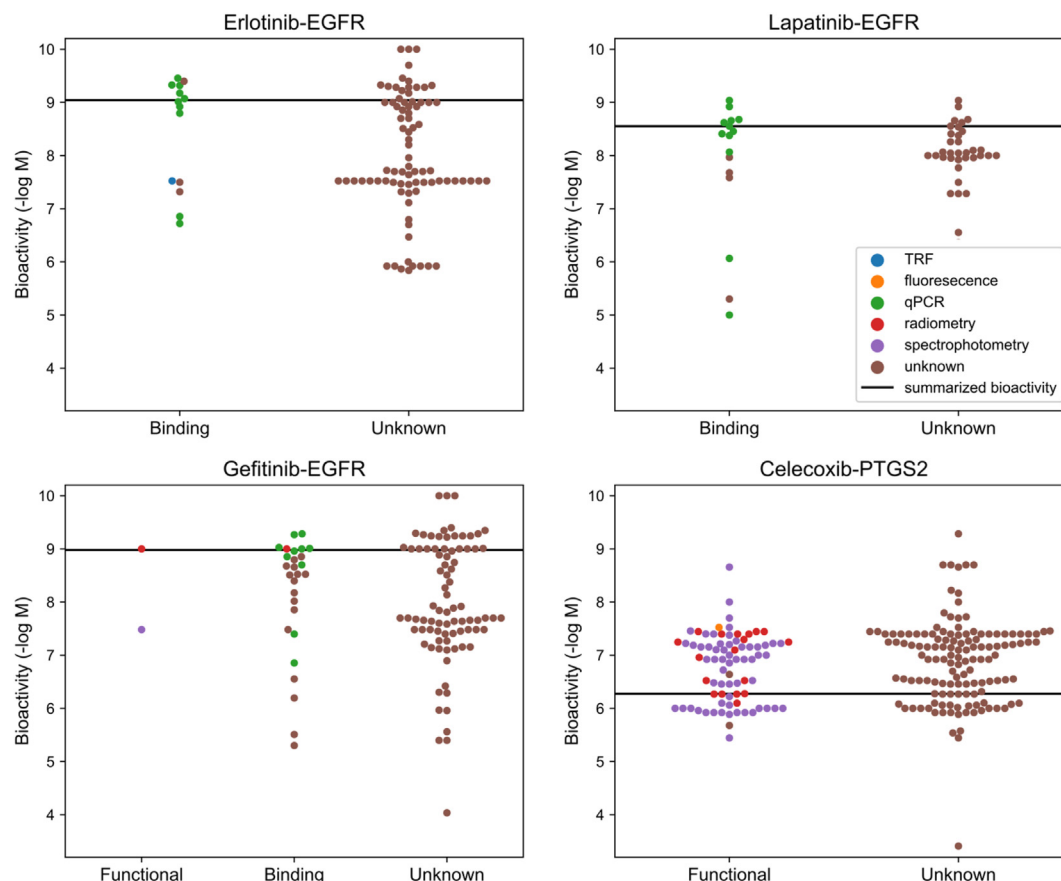


Fig. 2. Original and summarized bioactivities for four example drug-target pairs, where various assay types for the same pair are separated on the x-axis and the detection technologies used in the target profiling studies are indicated with colour coding of the bioactivity measurements (points). Higher values on y-axis indicate more potent compound-target binding affinities. TFR, Time Resolved Fluorescence; qPCR, quantitative Polymerase Chain Reaction.

than 100 potent targets (Fig. 4D). This demonstrates a rather wide spectrum of polypharmacological effects for many of the compounds in our library, consisting mostly of promiscuous kinase inhibitors (Fig. S1), which are known for their promiscuity and therefore require a multi-target deconvolution approach.

To obtain a more holistic view of the distribution of compound-target profiles in the dataset, the screened drug-target space was visualized with unsupervised t-Distributed Stochastic Neighbouring Embedding (t-SNE) mapping and hierarchical clustering (Fig. 5). Fig. 5A visualizes the underlying space of the screened drugs spanned by their binary target profile similarities (see Materials and Methods). Color-coding indicates the structural Tanimoto similarities of the nearest neighbour of each compound calculated using the MACCS fingerprints. The drug mapping reveals that many of the compounds with similar target profiles are also structurally similar (e.g., selumetinib and binimetinib). Fig. 5B shows the quantitative compound-target bioactivity matrix, where the protein targets are color-coded based on the protein superfamilies. This systematic analysis illustrates how the established drug and target classes are not very well clustered together in terms of the drug-target bioactivities, thus indicating that the bioactivity data provide additional information on the drugs' mode of action, not captured by their known mechanistic classes.

2.3. Functional dependencies and druggable vulnerabilities of TNBC cell lines

To quantify the functional importance of protein targets for the drug-cell line responses, TAS levels were calculated for each of the 632 targets, as the overall cell inhibition of compounds targeting

the particular protein, separately for the 20 TNBC cell lines, based on the summarized and filtered drug-target bioactivities (see Materials and Methods). Importantly, a permutation testing procedure was implemented to assess the statistical significance of the observed TAS values by means of calculating empirical p-values based on permutation null distribution. The correlation of TAS values from the two cell assays were compared to that of the original DSS values in the individual TNBC cell lines (Fig. 6). Since many of the drug compounds in our library have multiple active targets, and most of the targeted drugs exhibit cytostatic effect, the DSS values are expected to show relatively poor correlation between the toxicity and viability assays (Fig. 6A). However, in selected cell lines (e.g., CAL-51, DU4475, HCC1395), the TAS profiles based on the viability and toxicity assays showed correlated patterns (Fig. 6B), suggesting the TAS profiles were able to capture the common phenotypic changes in these cell line responses based on the two cell-based assays.

To further investigate how the bioactivity-based target profiles contribute to cell-based viability and toxicity responses of the drugs, we took example compound pairs from Fig. 5A, and further analysed their TAS and DSS patterns across the TNBC cell lines. In the first example of two drugs, luminespib and BIIB021, have rather dissimilar molecular structures, but they still showed similar target profiles (Fig. 7A). Notably, their common targets, HSP90AB1, HSP90B1, HSP90AA1 and TRAP1, were identified as the key dependencies in multiple cell lines (Fig. 7B), which accurately explained the drug response patterns across the heterogeneous TNBC cell lines (Fig. 7C). The second example shows how two structurally similar drugs, selumetinib and binimetinib

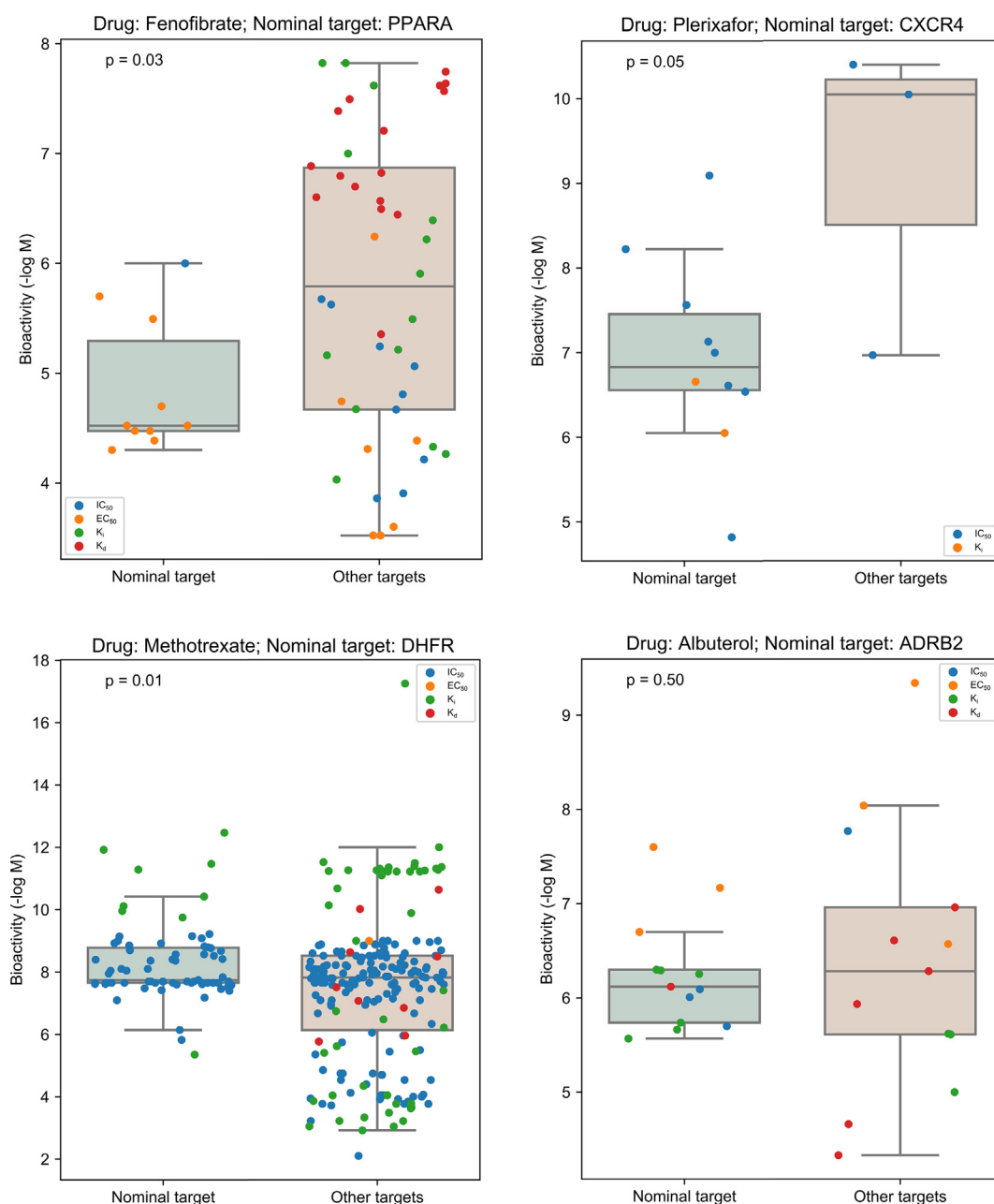


Fig. 3. Comparison of bioactivities of nominal targets and other targets for four example drugs: fenofibrate (nominal target: PPARA, Peroxisome proliferator-activated receptor alpha), plerixafor (nominal target: CXCR4, C-X-C motif chemokine receptor type 4), methotrexate (nominal target: DHFR, Dihydrofolate reductase), and albuterol (nominal target: ADRB2, Adrenoceptor Beta 2). The points in the boxplots indicate the dose-response bioactivity measurements before summarization of data from multiple target profiling studies and replicate measurements in DTC. The bioactivity distributions of nominal and other targets were compared by Mann-Whitney *U* test.

(Fig. S2A), which were grouped together in the t-SNE visualization (Fig. 5A), presented also with very similar TAS profiles for their two common targets, MAP2K1 and MAP2K2 (Fig. S2B). These examples demonstrate how the functional similarity based on the target dependencies provides further insights into the cell line-specific mechanisms of action of drugs, something that cannot be inferred from the structural similarity alone, or from the cell context-agnostic target profiles without using the TAS calculations.

2.4. Identification of key targets in each TNBC cell line based on drug and CRISPR screens

To identify context-specific therapeutic targets for TNBC, we calculated TAS values and empirical p-values that identified cell line-specific protein targets across the 20 TNBC cell lines. Targets

having a high TAS, along with a small empirical p-value, were identified as functionally important targets for a given cell line. As expected, the cell line-specific TAS profiles were able to identify a number of well-established TNBC driver genes (Table 2), such as AKT1 and AKT2 for CAL-148 and MFM-233; MAP2K1, MAP2K2, BIRC2, XIAP for MDA-MB-231 [26]. Furthermore, targeting the HDAC family and proteasome subunits (PSMB) tend to inhibit most of the TNBC cell lines, whereas CDK families were mostly specific to HCC1935. We also identified a number of so-far unexplored functionally important targets, such as STK11 for HCC-1395, TACR2 for BT-20 and PDXK for MFM-223 (Fig. 8; see also Fig. 6B), which warrant further studies by orthogonal cell assays. We note that some of the established drivers of the TNBC cell lines were not identified using the drug-induced TAS profiles (Table 2), which is

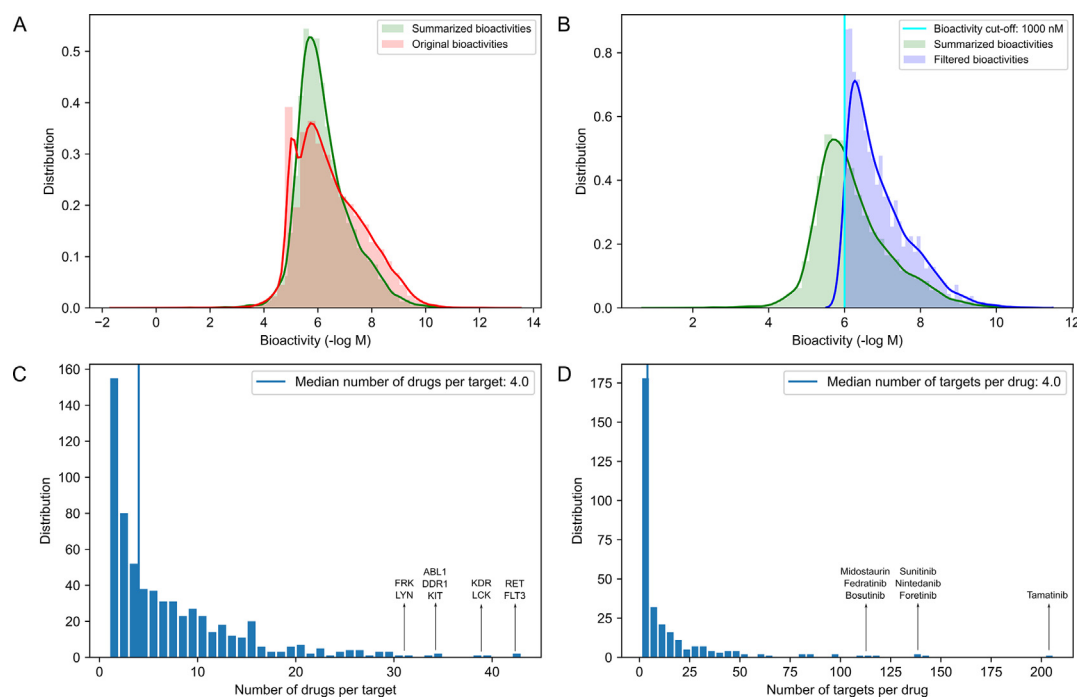


Fig. 4. A. Distribution of bioactivities before ($n = 23,469$) and after ($n = 8383$) the summarization process between 310 drugs and 782 targets. B. Distribution of bioactivities before ($n = 8383$, 310 drugs and 782 targets) and after ($n = 4276$, 310 drugs and 632 targets) the activity filtering (<1000 nM) process. C. Distribution of the number of potent drugs inhibiting each target after the summarization process and applying a 1000 nM activity cut-off for filtering, which led to a total of 632 potent targets of the 310 drugs; D. Distribution of the number of targets per drug after the bioactivity summarization and activity filtering process. Bioactivities larger than 1000 nM were regarded as non-potent, and were therefore not used in the further target deconvolution and other downstream analyses. The peak in the distribution of panel A corresponds to many 10000 nM values used in kinome target screens to represent non-active pairs.

likely due to their contribution to a cancer driving phenotype other than cell viability or death captured by our drug sensitivity assays. Notably, 56% of the functionally significant targets showed also elevated expression in the basal-like breast cancer patient samples (Fig. S3).

To make a more systematic analysis of the functional target dependency levels, TAS results were further corroborated by gene dependency scores derived from CRISPR-Cas9 gene knockout experiments in the same cell lines [27,28]. If a single gene knockout leads to reduced cell viability or cell death, then such gene is called as essential for the particular cell line, and thus has a more negative gene dependency score. Fig. 9 compares the TAS and gene dependency scores for each of the 632 targets in two example cell lines, BT-549 and HCC-1143. Surprisingly, many protein targets that had high TAS were associated with relatively high gene dependency scores, indicating that these targets are functionally essential in the cell context-specific drug mode-of-action, but not essential in terms of genetic knockout of the single genes. In fact, most of the protein targets with statistically significant TAS levels had close to zero gene dependency scores. This implies that these proteins may function as a group (or complex) when contributing to drug-cell line responses, since single knockout of any of these proteins shows only a limited impact on cell viability. We thus hypothesized that these functionally important protein groups that are resistant to single-gene knockout may pinpoint potential multi-target synthetic lethal interactions for the specific cell lines.

For the identification of multi-protein target groups, we explored how to make use of the pharmacological target dependency profiles of single drugs that modulate the function of several proteins simultaneously to exert multi-target co-inhibition effect. Such polypharmacological effects of promiscuous drugs are captured by the potent off-target interactions used in the TAS calculation, while being mostly missed in the single-gene CRISPR-Cas9

knockout screens. To more systematically study such multi-target interactions, we clustered the protein targets based on the potent drug-target interactions, with the aim to find protein target groups sharing common drugs as inhibitors. A number of protein groups were manually selected during the clustering process as their occurrence remains consistent across various parameters used in the t-SNE analyses (see Materials and Methods). One such protein group consisted of histone deacetylase (HDAC) enzymes (HDAC1, HDAC2, HDAC3, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9, HDAC10 and HDAC11, see Fig. S4). These enzymes are known to function in concert to remove acetyl groups from an amino acid on a histone, and may therefore jointly contribute to cell viability and selective killing. Based on these results, pan-HDAC inhibitors warrant further testing as potential multi-target co-inhibitors of HDAC synthetic lethal partners toward effective and stratified TNBC treatments.

3. Conclusions

We have developed a computational target deconvolution pipeline that implements several analysis tools for target dependency profiling, ranging from reliable construction of drug-target interaction networks to statistical analysis of the observed target addition levels. We demonstrated the operation of the pipeline in the context of highly heterogeneous TNBC cell lines that require cell line-specific analysis of drug and target vulnerabilities. The pipeline is expected to prove of value in multiple precision medicine applications also beyond target dependency profiling. For instance, comprehensive and confident compound-target interactions should prove useful also for predicting the phenotypic effects of combinatorial inhibitions [29,30], as well as to help manage potential side effects of combinatorial therapies [31].

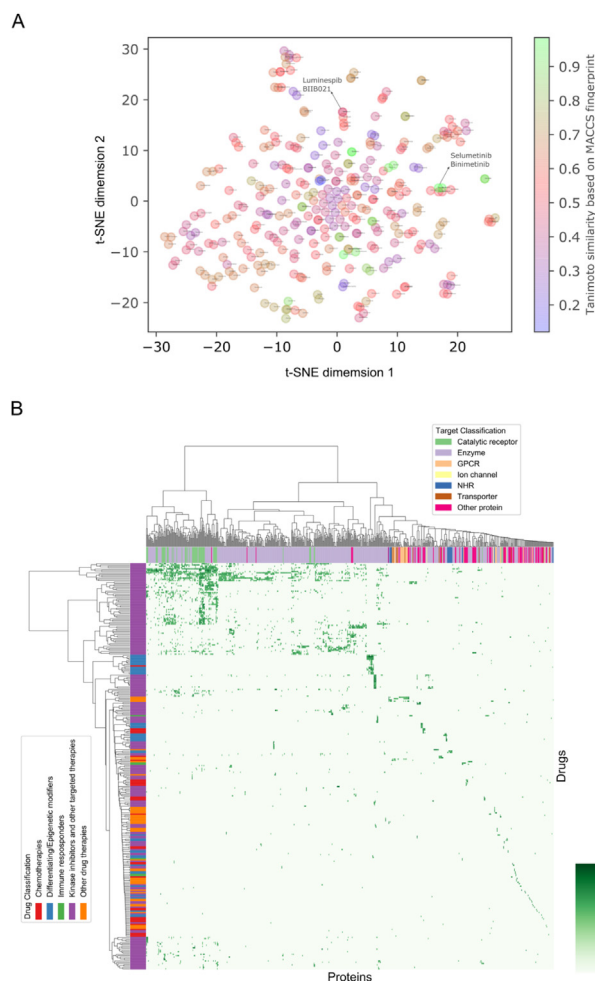


Fig. 5. A. t-SNE mapping of the 310 drugs based on the similarity of their binary target profiles, where drugs are color-coded by structural similarities of their neighbour compounds. B. Quantitative drug-target bioactivity matrix between 310 drugs and 632 targets, where the broad target and drug classes are marked with color-coding. NHR, Nuclear hormone receptors; GPCR, G protein-coupled receptors. The matrix includes the 4 276 unique bioactivities after bioactivity summarization and activity filtering, which covers 2% of the full drug-target matrix (the remaining entries of the matrix are either non-measured pairs, based on DTC data, or non-potent pairs, that is, summarized bioactivity >1000 nM).

Our target deconvolution pipeline makes use of various types of target bioactivity assays, both in cells and tissues, which are based on either biochemical or cell/tissue-based assay formats. The biochemical and cell-free assays are relatively straightforward in practice and can inform us about the specificity of compound-target interaction, whereas cell-based or physiochemical assays provide also physiological information of the compound-target binding. Furthermore, cell-based compound sensitivity assays capture the cell-specific characteristics of the compounds, such as their efficacy, toxicity and other cellular phenotypes, but these cannot quantify so accurately the specificity or selectivity of the compound-target interaction in the host cells. All these assays have factors that lead to variability and affect the consistency of target activity and cellular outcomes. In biochemical assays, for instance, factors such as pH, temperature, instrumentation, reagent characteristics and order of addition plays a vital role. Similarly, in cell-based assays, culture conditions such as humidity, temperature, culture media, serum, culture vessel, passage number, and cell cycle play a pivotal role.

There are also some notable differences between the 2D and 3D compound sensitivity assay formats. For instance, the 2D system

has a homogeneous population of proliferating cells grown in monolayer or suspension, and thereby receiving uniform supply of nutrients or exposure to the compounds. In contrast, 3D system comprises of cells in various stages, including proliferating, quiescent, apoptotic, necrotic and hypoxic stages, as well as different architectures, which results in rather different responses towards the compound treatment as compared to the 2D assays. Both assay outcomes also depend on the chosen time period of drug exposure (72 h in the present study). The more complex 3D systems, such as co-cultures or tissue slices, capture even more heterogeneous cell types, cell-cell and cell-matrix interactions, which also affect the compound responses. However, the main limitation of the 3D culture systems is that they are currently not compatible with high-throughput screening, which is important for systematic target deconvolution approaches and the main focus of this work.

Currently, there are not comprehensive enough datasets from any single assay type to cover all the relevant compounds and targets for systematic target deconvolution. Rather, compound-target bioactivity data are scattered in various databases, such as DTC and ChEMBL, that collect bioactivity data reported by different labs and studies using various experimental conditions. Therefore, in the absence of the knowledge of experimental factors affecting the assay outcomes, we performed in this work statistical consistency analysis to compare the various assay types in terms of their reproducibility and confidence (Table 1). This analysis can be readily extended once new assay types become available. This is one of the benefits of our generic target deconvolution pipeline that is applicable to any target profiling and compound sensitivity assays. We believe that accurate target identifications from the simpler experimental systems are critical for making translational discoveries for *in vivo* validations in animal and clinical studies in the future.

Empirical p-value calculation based on permutation testing was shown here to provide important information on the statistical support for an observed TAS level, e.g., when comparing TAS values that are based on different numbers of potent drugs (Fig. 4C), or across cell lines that show large variability in their sensitivity to the drugs (Fig. 7C). The robustness of the empirical p-value calculation to different drug library sizes was further investigated by subsampling drugs from the full library of 310 compounds. The permutation procedure was carried out by randomizing the drug-target links, while keeping the number of drugs fixed, in order to test how many drugs per target is enough to produce a reliable significance calculation. Comparing the p-value distribution from the subsampled drug sets against the original p-value from the full drug library indicated that the p-value calculation remains relatively robust until as few as 20% of the drugs are used in the target deconvolution (Fig. S5). Notably, already 2 active drugs per target was enough to provide rigid null distributions for the observed TAS values (Fig. S6). Smaller drug sets lead to fewer active drugs inhibiting some of the targets, and therefore the TAS values become increasingly dependent on a single or a few DSS values. For protein targets with only 1 potent drug, the distribution of the permuted TAS values closely resembles the distribution of DSS values of the same cell line (Fig. S6), as expected, indicating that such sparse drug-target interaction network in which each compound is connected to single protein cannot provide enough information for a reliable target deconvolution. We therefore recommend that for protein targets with permuted TAS values showing a close to uniform distribution, the observed TAS value is not reliable even if it has a significant p-value. With more potent drugs, there is enough statistical variability for empirical p-value calculation (Fig. S6).

Our target deconvolution method successfully recovered several well-known drivers of TNBC cell lines. Since HDACs regulate gene expression and transcription, many HDAC inhibitors, such

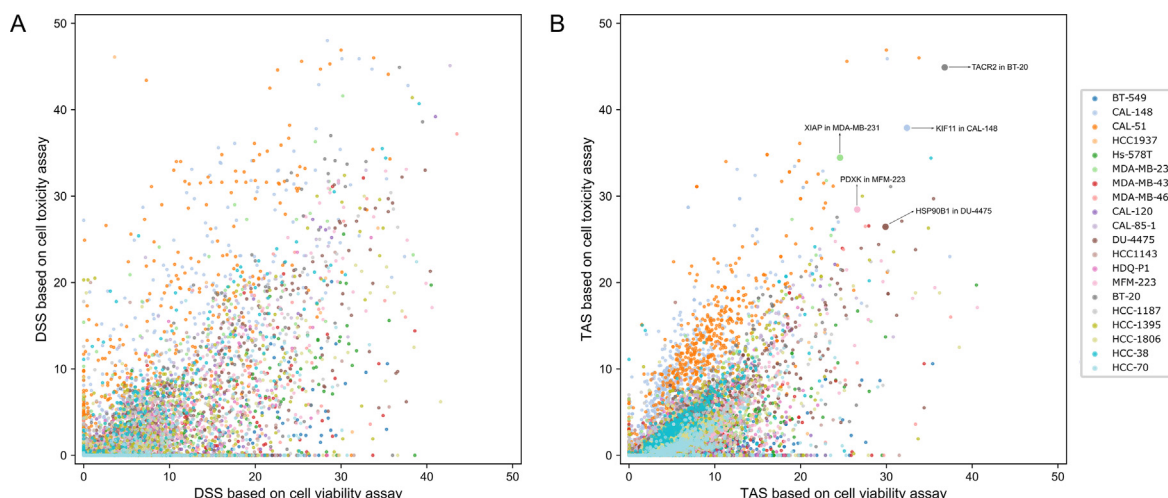


Fig. 6. A, B. Scatter plots of drug sensitivity score (DSS) of the 310 drugs (points in panel A) and target addiction score (TAS) of the 632 targets (panel B) between cell viability and cell toxicity assays calculated for the 20 TNBC cell lines (colours of the points). A higher DSS indicates increased sensitivity of a cell line to the particular drug, whereas a higher TAS suggests a higher functional dependency of the cell line on a specific target. Selected targets are highlighted in panel B as known drivers or potential key targets of the particular cancer cell lines with significant TAS value ($p < 0.05$).

as panobinostat and entinostat, have shown therapeutic effects in various cancer models and are currently undergoing clinical trials (e.g., NCT03361800) for breast cancer treatment [32]. Similarly, heat shock protein 90 (HSP90) is a protein class involved in regulating stability, activation and maturation of effector molecules driving many key signalling pathways, and hence explored as a target against TNBC since it disrupts multiple oncogenic pathways [33,34]. CDKs are key regulators of cell cycle, and TNBCs have been shown to be sensitive to CDK inhibitors [35,36]. For instance, CDK4/6 are regarded as promising targets for TNBC, and thereby extensively explored in the field [36–38]. We showed that TNBC cell lines, such as HCC38 and HCC70, have a unique dependence on CDK7, which can be harnessed as a selective therapeutic target [39]. Furthermore, inhibition of CDK12 and CDK13 provides synergistic effect on DNA damaging drugs or Poly (ADP-ribose) polymerase (PARP) inhibitors to induce TNBC cell death [40]. Inhibition of X-linked inhibitor of apoptosis (XIAP) in MDA-MB-231 has been shown to sensitize breast cancer cells to chemotherapies [41]. Aberrations in the phosphoinositide 3-kinase (PI3K)/protein kinase B (AKT)/mechanistic target of rapamycin (mTOR) pathway are also common in TNBC cell lines, and drugs targeting proteins in this pathway are well-studied and developed, such as PI3K and AKT inhibitors [42]. These and several other druggable targets, such as Kinesin Family Member 11 (KIF11) [43] and Aurora Kinase A (AURKA) [44], which were identified by the TAS analysis (Table 2), demonstrated its robustness to identify both known and novel functional addictions in individual cell lines.

Interestingly, the TAS analysis also identified several novel cell line-specific targets to achieve both cell viability inhibition and cell death. Some of the striking examples are tumour suppressor STK11 (upstream kinase of AMPK) for HCC-1395; a member of transmembrane G-protein coupled receptor TACR2, molecular motor protein KIF11 for BT-20, mitochondrial molecular chaperone TRAP1 (heat shock protein family) for DU4475; and kinase involved in vitaminB6 metabolism PDXK for MFM-223. Both BT-20 and MFM-223 were also identified to be affected by dual specificity protein kinase CLK3. Similarly, intestinal serine/threonine kinase ICK was identified to affect MFM-223, HCC-1395 and HCC-70. These so far unexplored target addictions warrant further studies. To understand the target pathway dependencies, we mapped 632 targets, pre-ranked by their TAS p-values from different cell lines, onto KEGG pathways using Gene set enrichment analysis (GSEA)

(Fig. S7). Significantly enriched pathways included AMPK signalling pathway, cell cycle, JAK-STAT signalling pathway, MAPK signalling pathway, p53 signalling pathway, viral carcinogenesis, inositol phosphate metabolism, and neuroactive ligand-receptor interaction ($FDR < 0.2$). Most of the effector molecules driving these pathways are known to be overexpressed in TNBC, making them potential therapeutic targets for TNBC. For example, AMPKs are relatively upregulated in TNBC as compared to non-TNBC [45], STAT3 is known to be overexpressed and constitutively active in TNBC [46], and reports show higher expression levels of MAPK in TNBC [47]. Hence, the significant target pathways further corroborated that targets with significant TAS levels are intensively involved in cell growth, apoptosis, and carcinogenesis. The main reason why the alcoholism pathway was significant was due to the enriched HDACs (Figure S7).

As a potential limitation, the current TAS profiling was not able to identify NTRK3/FGFR2 in MFM-223 and MAP2K1/MAP2K2 in MDA-MB-231 (Table 2). This is due to the general lack of selective drugs against these targets, or because the particular cell line was not sensitive to the potent drugs on average. For example, NTRK3 can be inhibited by 14 potent drugs in our library, with TAS of 4.66 and 1.17 from the cell viability and toxicity assays, respectively. However, the average TAS in MFM-223, calculated based on the 14 drugs in randomized drug-target interaction networks, are higher (6.93 and 2.77 from the two assays, respectively), which indicates that NTRK3 does not show statistically significant target dependency in the current dataset. Similarly, our algorithm did not capture CDK4/6 targets in the TNBC cell lines. This is because CDK4/6 inhibitors in our screening drug library did not exhibit strong effects on any of the TNBC cell lines in our panel. With larger drug libraries and more potent drugs, it is expected that also these established TNBC drivers could be identified using the target dependency profiling using solely the drug screening assays. Alternatively, one could combine TAS profiling with other cell line-specific experiments, such as gene expression or protein activity profiles, to make use of the baseline molecular background of the cell lines profiled before drug treatments. However, even if such additional profiling datasets can add another layer of context-specific information into the target inhibition analyses, these methods also add further costs to the studies, and especially with limited patient primary cells, one needs to carefully consider which assays can be afforded and be applied within a clinically

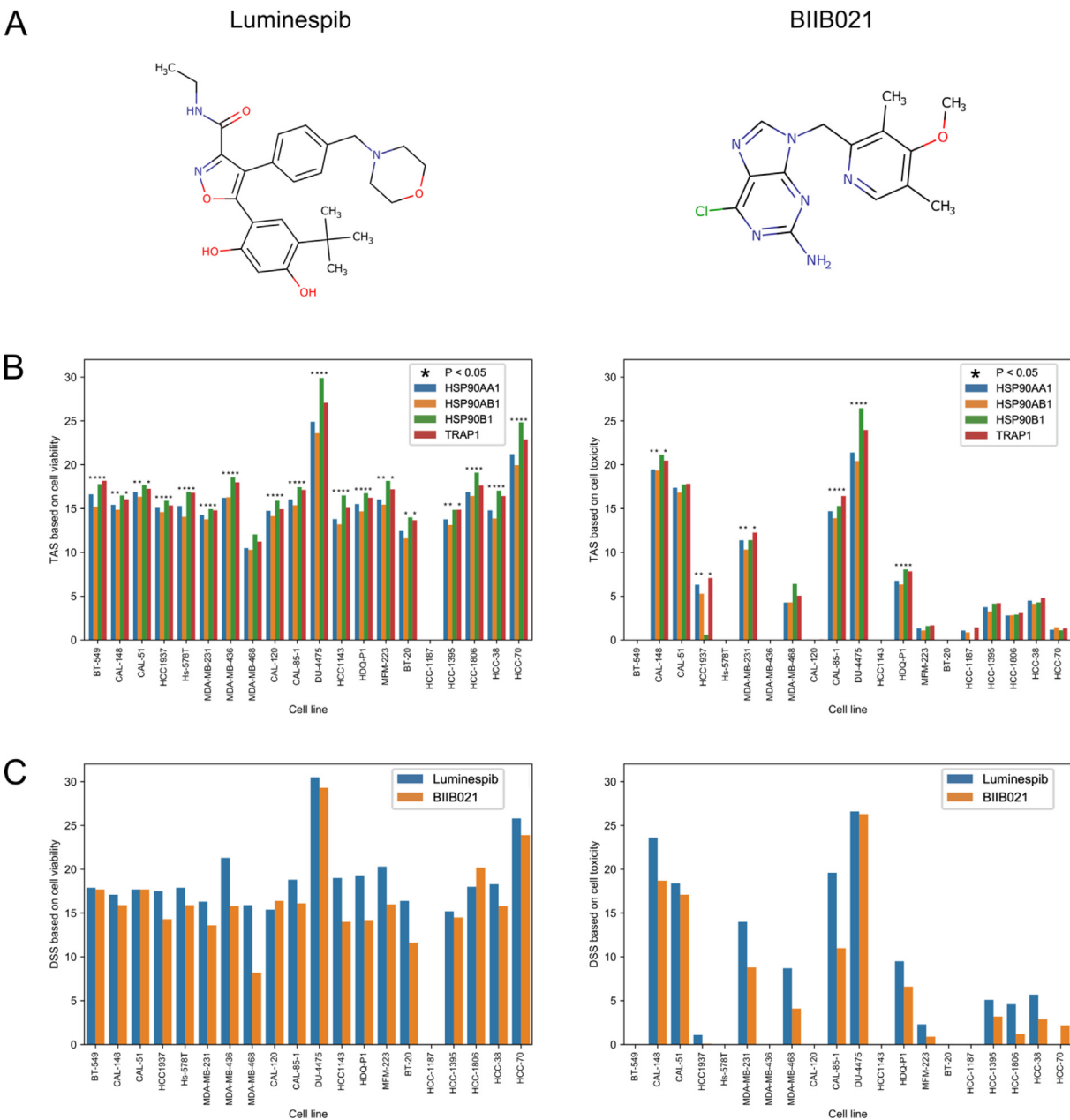


Fig. 7. A. 2D structures of luminespib and BIIB021 with low Tanimoto similarity of 0.452; B. TAS of their common targets, HSP90AB1, HSP90B1, HSP90AA1 and TRAP1, based on cell viability (left) and cell toxicity (right) assays; C. DSS of the drugs based on the viability and toxicity assays.

Table 2
Established drivers of TNBC cell lines and their TAS values calculated using cell viability and toxicity assays. Boldfacing indicates significant TAS values based on the permutation testing.

| Uniprot ID | Gene symbol | Cell line | TAS viability | P-value | Viability rank | TAS toxicity | P-value | Toxicity rank |
|------------|-------------|------------|---------------|---------------|----------------|--------------|---------------|---------------|
| P31749 | AKT1 | CAL-148 | 14,55 | 0,0050 | 14 | 17,77 | 0,0067 | 13 |
| P31751 | AKT2 | CAL-148 | 17,90 | 0,0040 | 13 | 22,44 | 0,0041 | 8 |
| P31749 | AKT1 | MFM-223 | 12,59 | 0,034 | 65 | 1,94 | 0,57 | 344 |
| P31751 | AKT2 | MFM-223 | 15,38 | 0,022 | 48 | 2,76 | 0,40 | 246 |
| P50613 | CDK7 | HCC-38 | 11,54 | 0,0014 | 5 | 6,39 | 0,022 | 26 |
| P50613 | CDK7 | HCC-70 | 11,69 | 0,0028 | 9 | 3,42 | 0,0084 | 22 |
| P98170 | XIAP | MDA-MB-231 | 24,60 | 0,0009 | 7 | 34,45 | 0,0000 | 1 |
| Q13490 | BIRC2 | MDA-MB-231 | 23,00 | 0,0025 | 14 | 31,80 | 0,0002 | 4 |
| P15056 | BRAF | DU-4475 | 13,57 | 0,055 | 62 | 10,84 | 0,011 | 22 |
| Q16288 | NTRK3 | MFM-223 | 4,66 | 0,86 | 578 | 1,17 | 0,86 | 463 |
| P21802 | FGFR2 | MFM-223 | 9,25 | 0,10 | 150 | 1,96 | 0,70 | 411 |
| Q02750 | MAP2K1 | MDA-MB-231 | 6,58 | 0,22 | 180 | 3,71 | 0,31 | 97 |
| P36507 | MAP2K2 | MDA-MB-231 | 6,44 | 0,28 | 220 | 5,13 | 0,17 | 51 |

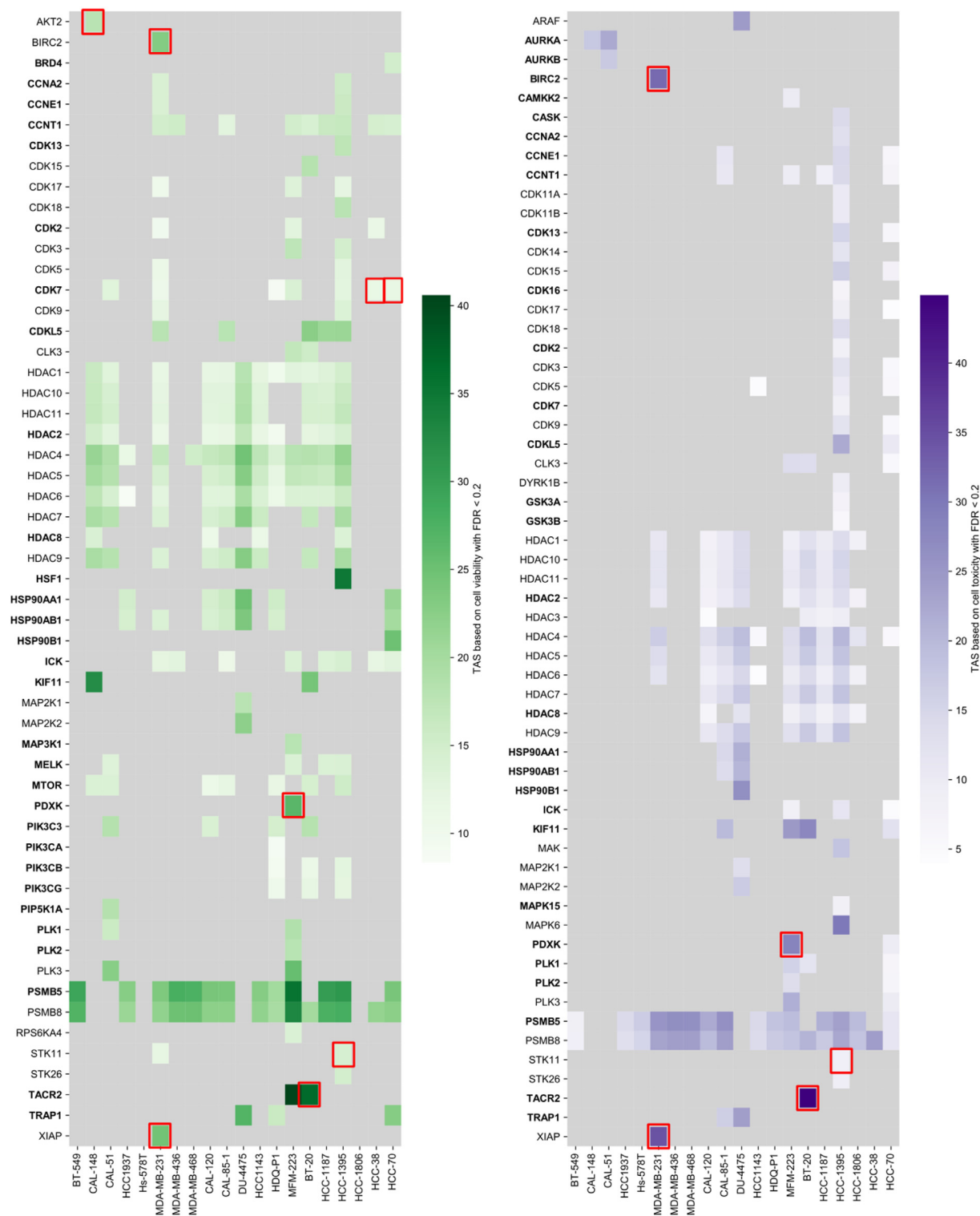


Fig. 8. Significant functionally important targets for each cell line (FDR < 0.2) based on cell viability (left) and toxicity (right) assays. Grey entries indicate non-significant cell line-target pairs (FDR ≥ 0.2). Boldfacing indicates that the protein targets discovered in the cell line screens were overexpressed in the basal-like breast cancer patient samples from TCGA (see Fig. S3).

actionable timeframe. The drug screening assays are fast and practical, taking 3 days to profile and 1 additional day to analyse the results.

There are many potential reasons why the target dependencies discovered based on the 2D drug screening data in TNBC cell lines do not necessary show as prognostic markers in breast cancer

patient tumors (Fig. S8). First, like noted above, there are fundamental differences between the *in vitro* and *in vivo* treatment responses. Second, target dependency is a cellular phenotype quantified after the *in vitro* treatments, whereas gene expression levels were measured in TCGA at baseline (before the patient treatment or after standard chemotherapy). Even though TCGA is the

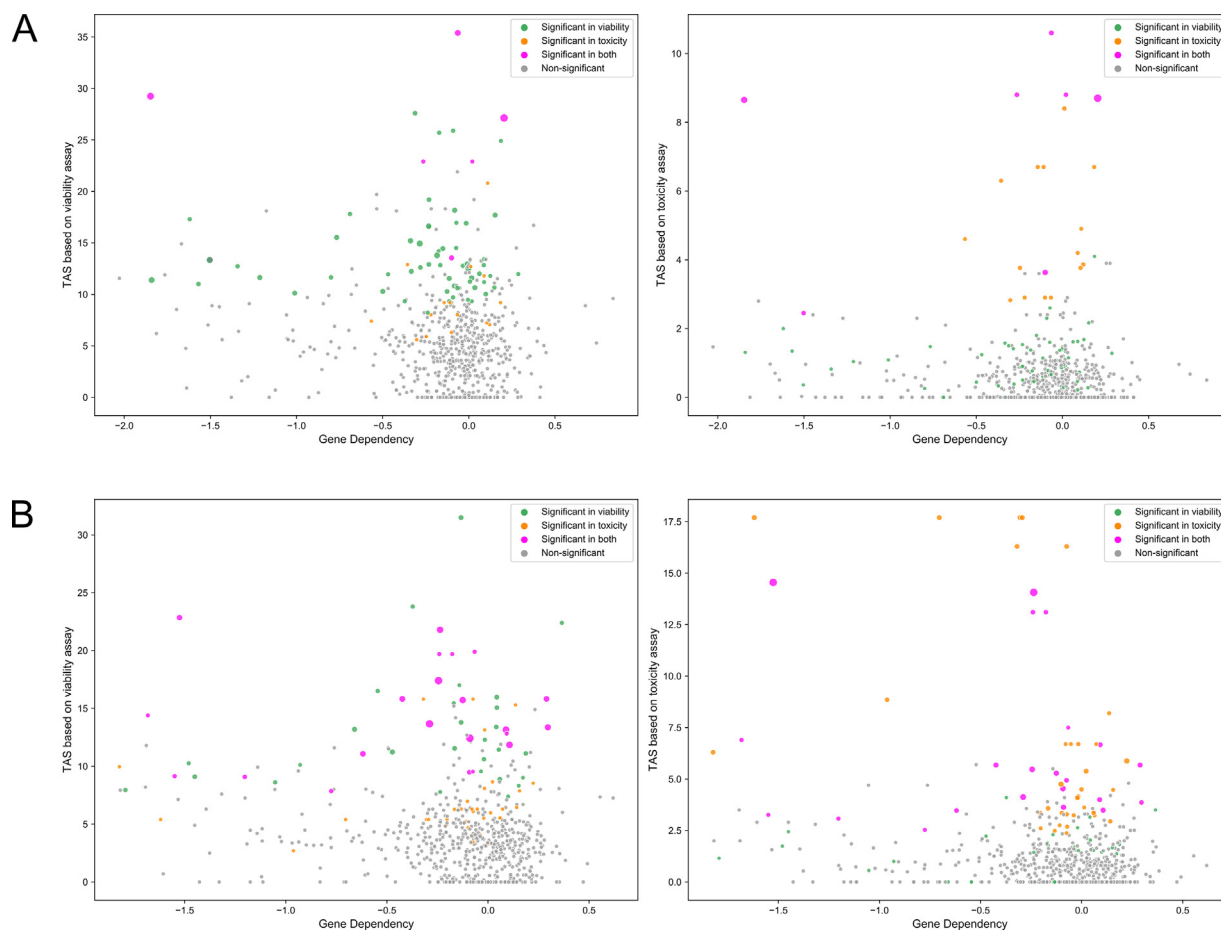


Fig. 9. A, B. Association between drug-induced functional dependency score (TAS) and gene dependency score from CRISPR-Cas9 screens across the 632 targets in BT-549 and HCC-1143 cell lines. The significance of the TAS levels based on the two assays is shown in color-coding, where empirical $p < 0.05$ indicates a statistically significant target (larger points).

largest cohort of cancer patients collected to date, it does not contain enough TNBC patients for statistically well-powered survival analyses (only 98 basal-like samples annotated as TNBC). Nevertheless, we found that 56% of the target additions discovered in the TNBC cell lines had a higher median gene expression level in the basal-like breast cancer samples (Fig. S3). This can be considered as relatively high statistical enrichment, given that our approach was not designed for the identification of prognostic protein markers for breast cancer patients, but rather to identify functionally important therapeutic targets based on cell line pharmacological screens. We believe the ideal way of using the target addiction analysis for cancer patients will be based on systematic drug screening in the patient-derived primary cells *ex-vivo*. The same target deconvolution approach can also be applied to patient-derived cells, as was earlier shown in the context of tailoring targeted treatments for patients with acute myeloid leukemia [48,49].

A recent study by Ann Lin et al. found out that off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials [11]. In particular, using single-target CRISPR-Cas9 assays, they demonstrated that a number of drugs exert their primary effects through other than their intended targets. This motivates the use of integrated approaches, such as TAS, which accounts for a wide spectrum of both weak and strong off-target effects of drugs in the target deconvolution analysis. We further demonstrated that single-gene dependency is not well-correlated with the TAS scores, suggesting that the CRISPR-Cas9 knockout

effect and the drug-induced inhibition effect are phenotypically quite different. Specifically, TAS provides information of druggable vulnerabilities of cell lines to a therapeutic targeting, as inferred from multiple drugs inhibiting the particular protein, whereas CRISPR-Cas9 provides information of the dependency of a cell line to the loss of the protein function. As shown in our results (Fig. 9), these two dependency scores provide complementary information for the cell context-specific targeting. We argue that TAS may be clinically more actionable since the druggability information is included in the analyses, hence offering pharmacologically more tractable cancer cell vulnerabilities for therapeutic targeting. For future precision oncology studies, integrated pharmacological and loss-of-function dependency analyses will likely offer improved possibilities to select optimal interventions [50], either functional or genetic perturbations, based on comprehensive information of both on/off targets of the drug treatment or CRISPR-Cas9-mediated genome engineering.

4. Materials & methods

4.1. Summarization of drug-target interactions based on various bioactivity measurements

The binding affinity of a drug against a protein target can be assessed with different experimental protocols and assays, and in various lab conditions. This results in multiple binding affinities for the same drug-protein pair, reported by different labs and stud-

ies, which hinders the integration of such bioactivity information into computational biology studies. To address the heterogeneity in reported bioactivity values, multiple binding affinity readouts for a given pair were summarized into a single harmonized value using our comprehensive DrugTargetCommons (DTC) database [17]. In DTC, part of bioactivities are fully annotated according to bioassay ontology (BAO), a simplified version of the minimum information about a bioactive entity (MIABE). With BAO, various bioassays conducted for the bioactivity assessments were annotated based on bioassay format, type and detection technology. Since the variability of replicated binding affinities from the same type of bioassay indicates the stability of this bioassay across different labs and studies, we used this categorization of bioassays to facilitate the summarization of various types of binding affinities for a drug-target pair.

Formally, let $D = \{d_i\}$ denote the set of all drugs; $T = \{t_j\}$ denote the set of all targets; $BT = \{bt_n\}$ denote the set of all bioactivity type categories; $AF = \{af_m\}$ denote the set of all assay format categories; $AT = \{at_l\}$ denote the set of all assay type categories; $DT = \{dt_k\}$ denote the set of all detection technology categories; and $B = \{b_{d_i, t_j, dt_k, at_l, af_m, bt_n} | d_i \in D, t_j \in T, dt_k \in DT, at_l \in AT, af_m \in AF, bt_n \in BT\}$ denote the set of all existing bioactivities.

The confidence ranking of a given assay condition (dt_k, at_l, af_m, bt_n) is formally calculated as

$$r_{dt_k, at_l, af_m, bt_n} = MAD(\{b_{d_i, t_j, dt_k, at_l, af_m, bt_n} | d_i \in D, t_j \in T\}) \quad (1)$$

Based on (1), the confidence ranking of the bioactivity types BT is defined as

$$R(bt_n) := Median(\{r_{dt_k, at_l, af_m, bt_n} | dt_k \in DT, at_l \in AT, af_m \in AF\}) \quad (2)$$

The summarized bioactivity between drug d_i and target t_j is then calculated as

$$bs_{d_i, t_j} = Median(\{b_{d_i, t_j, dt_k, at_l, af_m, bt_n} | bt_n = \underset{bt_n \in BT}{\operatorname{argmin}} R, dt_k \in DT, at_l \in AT, af_m \in AF\}) \quad (3)$$

In this study, four types of dose-response bioactivities, namely IC_{50} , EC_{50} , K_i and K_d , were extracted from the DTC database [17]. The confidence ranking of these bioactivity types was done separately for each bioassay. For a given bioactivity type, bioactivities were grouped by the assay condition. First, for each type of bioas-

say under each bioactivity type, bioactivities of all drug-target pairs were normalized to range [0,1]. Next, median absolute deviation (MAD) of multiple bioactivities was calculated as a variability estimate for each drug-target pair in each bioassay category. The median of all MADs within each bioassay type was used to evaluate the stability of the particular bioassay (Eq. (1)). The median of the median MADs across all bioassays within each bioactivity type was finally considered for ranking the various bioactivity types (Eq. (2)). Fig. 10 shows an example of calculating the ranking scores for the bioactivity type K_i , where the bioactivities were grouped based on assay type, assay format and detection technology. For those bioactivities in DTC with annotated bioassay information, the ranking of bioactivity type - assay format category was used to rank all the bioactivities of a drug-target pair. For those bioactivities with unknown bioassay information, the ranking of bioactivity type was used instead.

Finally, the multiple bioactivities of each drug-target pair were summarized into the median of bioactivities within the most confident ranking bioassay category (Eq. (3)). The summarized bioactivity values provide a standardized level of compound-target activity with the unit of nM, and these levels were used in the target deconvolution and other downstream analyses.

4.2. Calculation of drug sensitivity and target addiction scores in the TNBC cell lines

A panel of 20 breast cancer cell lines was screened against 310 oncology compounds in 5 different concentrations covering 10,000-fold concentration range using two cell health measurement assays, one for cell viability and the other for cell death, to differentiate between cytostatic and cytotoxic responses. For the cell viability measurements in the 2D cell-based *in vitro* assays, we used the luciferase-based cellular ATP detection reagent CellTiter-Glo (Promega); and for the cell death detection, the cell impermeable DNA-binding dye CellTox Green (Promega) was used [25]. Fig. S1 shows the distribution of the 310 approved drugs and investigational compounds, and Table S1 lists the 20 TNBC cell lines.

The drug sensitivity score (DSS) was calculated to quantify both the potency and the efficacy of the drug sensitivity of a cell line to a compound treatment. DSS was calculated for each cell line and compound pairs separately as the integral of the dose-response curve, which is modelled by logistic function through drug-cell responses under multiple doses [13]. For a given protein target,

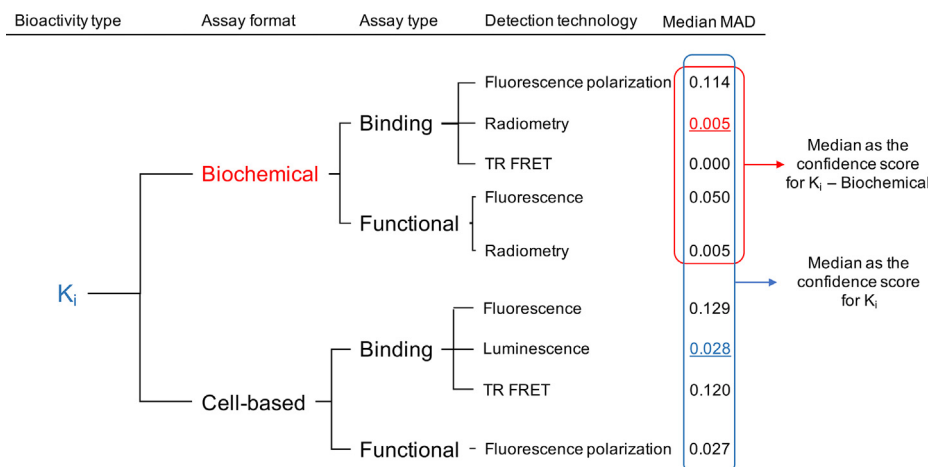


Fig. 10. Example of calculation of ranking score for the bioactivity type K_i . Median absolute deviation (MAD) of all drug-target pairs was first calculated within each category (e.g., K_i -Biochemical-Binding-Radiometry). Then, the median was used as the confidence score for this group. Furthermore, the confidence score for K_i -Biochemical is the median of all scores under this category. Confidence score of the bioactivity category K_i was calculated the same way.

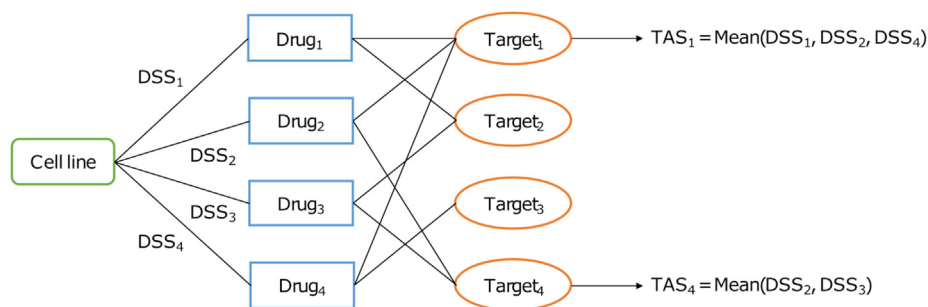


Fig. 11. Illustration of the calculation of target addiction score (TAS) for each protein target.

the target addiction score (TAS) was calculated as the average DSS of compounds known to target the protein (illustrated in Fig. 11), using 1000 nM as an activity cut-off value for the summarized compound–target bioactivities. TAS estimates the sensitivity of a cell line to the pharmacological inhibition of the protein target, as a measure of functional dependency of the cell line on the therapeutic target. DSS and TAS provide druggable insights into vulnerabilities of individual cells to therapeutic treatments and targets and were calculated as described before [13,17].

4.3. Identification of key protein target and protein groups in the TNBC cell lines

In the present work, we estimated the statistical significance of each observed TAS value by calculating empirical p-value using a permutation test. This is important, since the observed TAS values may not be directly comparable between studies, where different compound libraries are used, or between cell lines that have different compound response profiles. In contrast, p-values take into account such between study or cell line differences. Permutation tests for TAS values were carried out by randomizing compound–target interactions, while keeping the number of interactions for each drug unchanged. Permutation testing makes it possible to compare response profiles from drug testing experiments with different sizes of compound libraries. Targets having a high TAS, along with a small empirical p-value calculated based on the permutation null distribution, are identified as functionally important targets for a particular cell line. Subsampling of compounds and targets was also performed during the permutation test to find when the empirical p-value calculation becomes unreliable, as the number of compounds or targets decrease in the TAS calculation.

Publicly available gene essentiality scores based on CRISPR–Cas9 screens were obtained from the Dependency Map database (Release DepMap 19Q3) for the overlapping TNBC cell lines [27,28]. For each protein target, the available gene dependency score (CERES score) was matched by the HGNC official gene symbol and compared with the corresponding TAS value.

For identification of protein groups, a protein–drug interaction matrix was constructed, using summarized bioactivities from the DTC database. Each row of the matrix represents a drug bioactivity profile of a protein. If a protein–drug interaction potency is smaller than 1000 nM, it is encoded as “1”, otherwise as “0”. Using these binary drug profiles as features, all the proteins were mapped into 2D planes with the t-SNE method with L1 and L2 norms as distance metrics, respectively [51]. Proteins constantly appeared as groups across all the parameters of the t-SNE methods were manually picked as groups of interests. For specific cell lines, the key actionable protein groups were identified as those with more than 3 active drugs and significant TAS values ($p < 0.05$). Protein targets, pre-ranked by their p-values in each cell line, were mapped onto

KEGG pathways and Gene Ontology (GO) biological processes to provide further mechanistic information, using the gene set enrichment analysis (GSEA) [52], implemented through GSEapy python package (github.com/zqfang/GSEapy).

5. Funding acknowledgement

The work was partially funded by the Academy of Finland (grants 310507, 313267, 326238 to TA; 334790 to JR), Helse Sør-Øst (grant 2020026), Cancer Society of Finland (TA), Sigrid Jusélius Foundation (TA), and Doctoral Program in Integrative Life Science (TW).

CRediT authorship contribution statement

Tianduanyi Wang: Conceptualization, Data curation, Formal analysis, Methodology, Visualization, Writing - original draft. **Prson Gautam:** Validation, Investigation, Resources, Writing - review & editing. **Juho Rousu:** Funding acquisition, Supervision, Writing - review & editing. **Tero Aittokallio:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2020.11.001>.

References

- [1] Zheng W, Thorne N, McKew JC. Phenotypic screens as a renewed approach for drug discovery. *Drug Discov Today* 2013;18(21–22):1067–73.
- [2] Macarron R, Banks MN, Bojanic D, Burns DJ, Cirovic DA, Garyantes T, et al. Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* 2011;10(3):188–95.
- [3] Terstappen GC, Schlupen C, Raggiaschi R, Gaviraghi G. Target deconvolution strategies in drug discovery. *Nat Rev Drug Discov* 2007;6(11):891–903.
- [4] Kubota K, Funabashi M, Ogura Y. Target deconvolution from phenotype-based drug discovery by using chemical proteomics approaches. *Biochim Biophys Acta Proteins Proteom* 2019;1867(1):22–7.
- [5] Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. *Nat Rev Cancer* 2009;9(1):28–39.
- [6] Kumar R, Singh VP, Baker KM. Kinase inhibitors for cardiovascular disease. *J Mol Cell Cardiol* 2007;42(1):1–11.
- [7] Garon SL, Pavlos RK, White KD, Brown NJ, Stone Jr CA, Phillips EJ. Pharmacogenomics of off-target adverse drug reactions. *Br J Clin Pharmacol* 2017;83(9):1896–911.
- [8] Anighoro A, Bajorath J, Rastelli G. Polypharmacology: challenges and opportunities in drug discovery. *J Med Chem* 2014;57(19):7874–87.

- [9] Ravikumar B, Aittokallio T. Improving the efficacy-safety balance of polypharmacology in multi-target drug discovery. *Expert Opin Drug Discov* 2018;13(2):179–92.
- [10] Kaelin Jr WG. The concept of synthetic lethality in the context of anticancer therapy. *Nat Rev Cancer* 2005;5(9):689–98.
- [11] Lin A, Giuliano CJ, Palladino A, John KM, Abramowicz C, Yuan ML, et al. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Sci Transl Med* 2019;11(509). eaaw8412.
- [12] Hopkins AL. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* 2008;4(11):682–90.
- [13] Yadav B, Pemovska T, Szwajda A, Kuleskiy E, Kontro M, Karjalainen R, et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci Rep* 2014;4:5193.
- [14] Mpindi JP, Yadav B, Östling P, Gautam P, Malani D, Murumägi A, et al. Consistency in drug response profiling. *Nature* 2016;540(7631):E5–6.
- [15] Hafner M, Niepel M, Chung M, Sorger PK. Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. *Nat Methods* 2016;13(6):521–7.
- [16] Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, et al. ChEMBL: towards direct deposition of bioassay data. *Nucl Acids Res* 2019;47(D1):D930–40.
- [17] Tang J, Tanoli ZU, Ravikumar B, Alam Z, Rebane A, Vaha-Koskela M, et al. Drug target commons: a community effort to build a consensus knowledge base for drug-target interactions. *Cell Chem Biol* 2018;25(2). 224–9 e2.
- [18] Yadav B, Gopalacharyulu P, Pemovska T, Khan SA, Szwajda A, Tang J, et al. From drug response profiling to target addiction scoring in cancer cell models. *Dis Model Mech* 2015;8(10):1255–64.
- [19] Bareche Y, Venet D, Ignatiadis M, Aftimos P, Piccart M, Rothe F, et al. Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. *Ann Oncol* 2018;29(4):895–902.
- [20] Garrido-Castro AC, Lin NU, Polyak K. Insights into molecular classifications of triple-negative breast cancer: improving patient selection for treatment. *Cancer Discov* 2019;9(2):176–98.
- [21] Mehanna J, Haddad FG, Eid R, Lambertini M, Kourie HR. Triple-negative breast cancer: current perspective on the evolving therapeutic landscape. *Int J Womens Health* 2019;11:431–7.
- [22] Lee KL, Kuo YC, Ho YS, Huang YH. Triple-negative breast cancer: current understanding and future therapeutic breakthrough targeting cancer stemness. *Cancers (Basel)* 2019;11(9).
- [23] Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat Rev Clin Oncol* 2016;13(11):674–90.
- [24] Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol* 2008;26(8):1275–81.
- [25] Gautam P, Karhinen L, Szwajda A, Jha SK, Yadav B, Aittokallio T, et al. Identification of selective cytotoxic and synthetic lethal drug responses in triple negative breast cancer cells. *Mol Cancer* 2016;15(1):34.
- [26] Gautam P, Jaiswal A, Aittokallio T, Al-Ali H, Wennerberg K. Phenotypic screening combined with machine learning for efficient identification of breast cancer-selective therapeutic targets. *Cell Chem Biol* 2019;26(7):970–9 e4.
- [27] Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* 2017;49(12):1779–84.
- [28] Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, et al. Defining a cancer dependency map. *Cell* 2017;170(3):564–76 e16.
- [29] Menden MP, Wang D, Mason MJ, Szalai B, Bulusu KC, Guan Y, et al. Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. *Nat Commun* 2019;10(1):2674.
- [30] Tang J, Gautam P, Gupta A, He L, Timonen S, Akimov Y, et al. Network pharmacology modeling identifies synergistic Aurora B and ZAK interaction in triple-negative breast cancer. *npj Syst Biol Appl* 2019;5(1).
- [31] Giri AK, Ianevski A, Aittokallio T. Genome-wide off-targets of drugs: risks and opportunities. *Cell Biol Toxicol.* 2019;35(6):485–7.
- [32] Tate CR, Rhodes LV, Segar HC, Driver JL, Pounder FN, Burow ME, et al. Targeting triple-negative breast cancer cells with the histone deacetylase inhibitor panobinostat. *Breast Cancer Res* 2012;14(3):R79.
- [33] Nakhjavani M, Hardingham JE, Palethorpe HM, Price TJ, Townsend AR. Druggable molecular targets for the treatment of triple negative breast cancer. *J Breast Cancer* 2019;22(3):341–61.
- [34] Proia DA, Zhang C, Sequeira M, Jimenez JP, He S, Spector N, et al. Preclinical activity profile and therapeutic efficacy of the HSP90 inhibitor ganetespib in triple-negative breast cancer. *Clin Cancer Res* 2014;20(2):413–24.
- [35] Matutino A, Amaro C, Verma S. CDK4/6 inhibitors in breast cancer: beyond hormone receptor-positive HER2-negative disease. *Ther Adv Med Oncol* 2018;10. 1758835918818346.
- [36] Asghar US, Barr AR, Cutts R, Beaney M, Babina I, Sampath D, et al. Single-cell dynamics determines response to CDK4/6 inhibition in triple-negative breast cancer. *Clin Cancer Res* 2017;23(18):5561–72.
- [37] Ni Y, Schmidt KR, Werner BA, Koenig JK, Guldner IH, Schnepf PM, et al. Death effector domain-containing protein induces vulnerability to cell cycle inhibition in triple-negative breast cancer. *Nat Commun* 2019;10(1):2860.
- [38] Dai M, Zhang C, Ali A, Hong X, Tian J, Lo C, et al. CDK4 regulates cancer stemness and is a novel therapeutic target for triple-negative breast cancer. *Sci Rep* 2016;6(1):35383.
- [39] Wang Y, Zhang T, Kwiatkowski N, Abraham BJ, Lee TI, Xie S, et al. CDK7-dependent transcriptional addiction in triple-negative breast cancer. *Cell* 2015;163(1):174–86.
- [40] Quereda V, Bayle S, Vena F, Frydman SM, Monastyrskyi A, Roush WR, et al. Therapeutic targeting of CDK12/CDK13 in triple-negative breast cancer. *Cancer Cell* 2019;36(5):545–58 e7.
- [41] McManus DC, Lefebvre CA, Cherton-Horvat G, St-Jean M, Kandimalla ER, Agrawal S, et al. Loss of XIAP protein expression by RNAi and antisense approaches sensitizes cancer cells to functionally diverse chemotherapeutics. *Oncogene* 2004;23(49):8105–17.
- [42] Costa RLB, Han HS, Gradishar WJ. Targeting the PI3K/AKT/mTOR pathway in triple-negative breast cancer: a review. *Breast Cancer Res Treat* 2018;169(3):397–406.
- [43] Lucanus AJ, Yip GW. Kinesin superfamily: roles in breast cancer, patient prognosis and therapeutics. *Oncogene* 2018;37(7):833–8.
- [44] Chang SS, Yamaguchi H, Xia W, Lim SO, Khotskaya Y, Wu Y, et al. Aurora A kinase activates YAP signaling in triple-negative breast cancer. *Oncogene* 2017;36(9):1265–75.
- [45] Huang X, Li X, Xie X, Ye F, Chen B, Song C, et al. High expressions of LDHA and AMPK as prognostic biomarkers for breast cancer. *Breast* 2016;30:39–46.
- [46] Qin JJ, Yan L, Zhang J, Zhang WD. STAT3 as a potential therapeutic target in triple negative breast cancer: a systematic review. *J Exp Clin Cancer Res* 2019;38(1):195.
- [47] Jiang W, Wang X, Zhang C, Xue L, Yang L. Expression and clinical significance of MAPK and EGFR in triple-negative breast cancer. *Oncol Lett* 2020;19(3):1842–8.
- [48] Malani D, Murumagi A, Yadav B, Kontro M, Eldfors S, Kumar A, et al. Enhanced sensitivity to glucocorticoids in cytarabine-resistant AML. *Leukemia* 2017;31(5):1187–95.
- [49] Pemovska T, Kontro M, Yadav B, Edgren H, Eldfors S, Szwajda A, et al. Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer Discov* 2013;3(12):1416–29.
- [50] Gonçalves E, Segura-Cabrera A, Pacini C, Picco G, Behan FM, Jaaks P, et al. Drug mechanism-of-action discovery through the integration of pharmacological and CRISPR screens. *Mol Syst Biol* 2020;16(7):e9405.
- [51] Lvd M, Hinton G. Visualizing data using t-SNE. *J Machine Learn Res* 2008;9(Nov):2579–605.
- [52] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102(43):15545–50.